XXVI. PROCESSING AND TRANSMISSION OF INFORMATION

Academic and Research Staff

Prof. P. Elias	Prof. E. V. Hoversten	Prof. E. Mortenson
Prof. R. M. Gallager	Prof. D. A. Huffman	Prof. C. E. Shannon
Prof. M. M. Goutmann	Prof. R. S. Kennedy	Prof. R. N. Spann
Prof. F. C. Hennie III	Prof. J. L. Massey	Prof. J. T. Wagner

Graduate Students

D. S. Arnstein	M. Khanna	J. T. Pinkston III
E. A. Bucher	Jane W-S. Liu	E. M. Portner, Jr.
D. Chase	J. Max	J. S. Richters
R. L. Greenspan	J. C. Moldon	A. H. M. Ross
H. M. Heggestad	R. Pilc	S. Thongthammachat
J. A. Heller		D. A. Wright

RESEARCH OBJECTIVES AND SUMMARY OF RESEARCH

1. Communications

The work of this group is focused on the dual problems of ascertaining the best performance that can be attained with a communication system, and developing efficient techniques for actually achieving performances substantially this good.

a. Convolutional Codes and Decoding

A technique has been investigated for combining block coding with convolutional coding.[1] Using a combination of sequential decoding and algebraic decoding, we have demonstrated that reliable communication can be achieved at all rates below channel capacity. While this technique is more complicated than sequential decoding alone, it operates at higher rates and has less buffer storage requirements than sequential decoding.

Preliminary results have been achieved on the relative probability performance of nonsystematic versus systematic convolutional codes.[2] Surprisingly, the results indicate that for nonsystematic codes, the exponential decay of error probability with code constraint length is much faster than for systematic codes.

A theoretical investigation has been undertaken on the capabilities of tree codes for error-correction purposes, fixed convolutional codes being treated as an important special case. Minimum distances of an appropriate type have been defined for both feedback decoding and nonfeedback or definite decoding of tree codes. Some new upper and lower bounds on minimum distance have been obtained and effort in this direction continues. The error-propagation effect resulting from feedback decoding is also being investigated. A new decoding technique, called semidefinite decoding, which has some of the features of feedback decoding but avoids the error-propagation effect, is being simulated for a wide variety of convolutional codes on the binary symmetric channel to determine whether it offers an over-all advantage over feedback decoding.

J. L. Massey, R. G. Gallager

---

*This work was supported in part by the National Aeronautics and Space Administration (Grants NsG-334 and NsG-496); and in part by the Joint Services Electronics Programs (U.S. Army, U.S. Navy, and U.S. Air Force) under Contract DA 36-039-03200(E), the National Science Foundation (Grant GP-835), and the Sloan Fund for Basic Research (M.I.T. Grant).

References

1. D. Falconer, "A Hybrid Sequential and Algebriac Decoding Scheme," Ph.D. Thesis, Department of Electrical Engineering, M.I.T., September 1966.
2. E. A. Bucher, "Error Probability for Systematic Convolutional Codes," Thesis research, in preparation.

b. Optical Communication

The extension of communication theory to optical channels has been focused on the effects of atmospheric turbulence. A model that is appropriate for the analysis of communication systems has been developed with the techniques of geometric optics. Since our approach provides some insights into more general problems of atmospheric propagation, it is being developed beyond the level required for communication theory to its natural conclusion. The character of these conclusions is discussed in this report (see Sec. XXVI-B).

A 4-km one-way propagation path operating at 6328 Å has been established with the with the cooperation of the Harvard College Observatory.[1] The facility has been used to investigate the statistical properties of intensity, and is now being modified for round-trip transmission. An interferometric system has also been developed to study turbulence-induced phase fluctuations.[2] A theoretical and experimental study concerning depolarization caused by turbulence has led to the conclusion that depolarization is negligible.[3] The experimental aspects of the study were carried on at Bell Telephone Laboratories, Inc., Crawford Hill.

Recently initiated investigations include studies of the communication reliability of free space and atmospheric channels in the absence of quantum effects, the reliability of quantum channels, the potential of forward-scatter communication systems, and the estimation of incoherently illuminated objects viewed through a turbulent atmosphere.

R. S. Kennedy, E. V. Hoversten

References

1. J. E. Roberson, "A Study of Atmospheric Effects on Intensity Spatial and Temporal Properties at 6328 Å," S.M. Thesis, Department of Electrical Engineering, M.I.T., October 1966.
2. R. Yusek, "Interferometric Measurement of Optical Phase Noise over Atmospheric Paths," S.M. Thesis, Department of Electrical Engineering, M.I.T., October 1966.
3. A. A. M. Saleh, "Laser Wave Polarization by Atmospheric Transmission," S.M. Thesis, Department of Electrical Engineering, M.I.T., November 1966.

c. Specific Channels and Coding

One of the major types of channels which is being investigated is a class of fading dispersive channels such as HF and Tropo. Particular emphasis is being placed upon the situation in which the information rate is comparable to, or exceeds, the available bandwidth; the complementary situation has been considered elsewhere.[1]

The performance that can be achieved with one particular signaling scheme when

both the coherence bandwidth and the frequency dispersion of the channel are small has been investigated analytically and experimentally.[2] The results suggest that a very large energy-to-noise ratio is required for satisfactory communication when the rate is comparable to the bandwidth. A comprehensive analytical study of the attainable reliability is now in progress. Partial results are presented in this report (see Sec. XXVI-A).

A number of coding theorems have been developed for statistically related parallel channels in the absence of cross talk.[3] Such models can be applied, for example, to frequency-multiplexed channels and also yield insight into the behavior of channels with memory.

A theoretical investigation of the use of coding on unsynchronized, noisy channels is in progress.[4] The lack of synchronization does not change the random coding exponent, but appears to radically change the exponent to error probability at low rates or for channels with little noise.

R. S. Kennedy, R. G. Gallager

References

1. R. S. Kennedy, Fading Dispersive Communication Channels (John Wiley and Sons, Inc., in press).
2. J. Moldon, "High Rate Reliability over Fading Dispersive Communication Channels," S. M. Thesis, M. I. T., Department of Electrical Engineering, June 1966.
3. J. Max, "Parallel Channels without Cross Talk," Thesis research, in preparation.
4. D. Chase, "Communication over Noisy Channels with No a priori Synchronization Information," Thesis research, in preparation.

d. Coding and the Processing of Information

The problem of performing reliable computation with unreliable computing elements has been reviewed by Winograd and Cowan.[1] Earlier work in this field has led to results that enable improvement in the reliability of computation only by increasing the number of unreliable elements used per unit computation, or increasing the complexity of each such element without reducing its reliability. Results analogous to the noisy-channel coding theorem of information theory, which would permit increasing the reliability of computation while performing more of it at once, with a fixed redundancy of equipment per unit computation, have not been available. Michael C. Taylor[2] has shown that application of low-density parity-check codes[3] to the problem of reliable storage of information in a noisy register leads to a result of coding theorem character: Doubling both the number of nosiy components and the amount of stored information reduces the probability of error or, alternatively, increases the mean time until an error occurs.

Two other topics relating coding to information processing are now under investigation. The first is the efficient addressing of stored data. The second is the trade-off between informational efficiency in encoding the output of an analog source and the ability to answer a variety of questions from the encoded output, or to tolerate a range of possible source characteristics. The second question is of interest for telemetry and other analog-digital conversion applications.

P. Elias, R. Gallager

References

1. S. Winograd and J. C. Cowan, Reliable Computation in the Presence of Noise (The M. I. T. Press, Cambridge, Mass., 1963).
2. M. C. Taylor, "Randomly Perturbed Computer Systems," Ph.D. Thesis, Department of Electrical Engineering, M. I. T., September 1966.
3. R. G. Gallager, Low Density Parity Check Codes (The M. I. T. Press, Cambridge, Mass., 1963).

e. Networks of Nosiy Channels

Earlier results concerning particular two-terminal networks of channels perturbed by additive Gaussian noise[1] have been extended to a large class of such networks, including, among others, series-parallel and bridge networks. The best signal-to-noise ratio attainable at the output by taking linear combinations of signals arriving at a node has been determined for this class, and has been bounded for all two-terminal networks. This work has been presented orally[2] and submitted for publication. Further work on numerical evaluation of cases of feedback networks not included in the solved class is under way.

P. Elias

References

1. P. Elias, "Channel Capacity without Coding," Quarterly Progress Report, Research Laboratory of Electronics, M. I. T., October 15, 1956, pp. 90-93.
2. P. Elias, "Networks of Gaussian Channels with Application to Feedback Systems," presented at the International Congress of Mathematicians, Moscow, August 1966.

f. Source Coding with a Distortion Measure

The interrelations between source and channel coding have been investigated for discrete memoryless sources and channels.[1] The results indicate that for combined source and channel coding with block length n, the theoretical minimum distortion is approached with increasing n, the dependence of the rate of approach with n being between $1/n$ and $\sqrt{\ln n/n}$.

In another investigation, a distortion measure for a discrete source was considered with a distortion of 1 for error, and 0 for no error. This led to a complete solution for the minimum probability of error achievable for a discrete memoryless source when transmitting over a channel with capacity less than the source entropy.[2]

Further investigations are being made on techniques for source coding and on the effect of quantizers in source coding.

R. G. Gallager, R. S. Kennedy

References

1. R. Pilc, "Coding Theorems for Discrete Source-Channel Pairs," Ph.D Thesis, Department of Electrical Engineering, M.I.T., November 1966.
2. J. T. Pinkston, "An Application of Rate Distortion Theory to a Converse to the Coding Theorem" (submitted for publication to IEEE Transactions on Information Theory).

A. LOW-RATE UPPER BOUNDS ON ERROR PROBABILITY FOR FADING DISPERSIVE CHANNELS

1. Introduction

A scatter communication channel such as an orbital dipole belt can frequently be considered as a collection of K equal-strength diversity paths, each with independent additive Gaussian noise.[1] We consider using such a channel to transmit amplitude information, x, on some basic unit energy signal. Under these conditions, it is possible to define a statistic, y, at the channel output that is sufficient for the estimation of x, where

$$p_K(y|x) = \frac{y^{2K-1} e^{-\frac{y^2/2}{\frac{x^2}{K}+N_o}}}{2^{K-1}\Gamma(K)\left(\frac{x^2}{K}+N_o\right)^K} \tag{1}$$

and the noise power density is $N_o/2$ watts/cps. This abstraction results in a continuous input-continuous output channel with input x and output y governed by the probability function $p_K(y|x)$.

By transmitting time and frequency translates of the basic signal, and providing sufficient guard space in both time and frequency, we can obtain N independent channel uses. We consider transmission of one of $M = e^{NR}$ equally probable input signals and represent the m^{th} input signal as an N-tuple $\{x_{1m}, x_{2m}, \ldots, x_{Nm}\}$, where x_{nm} is the amplitude of the n^{th} channel use. Such a communication scheme allows us to explicity take into account a bandwidth constraint for communication over a fading dispersive channel, by relating N and K to W, the total input bandwidth, and T, the total time duration of the input signals.

2. Error Probability

Under the conditions just outlined above, we may apply the random coding upper bound to error probability, as discussed by Gallager.[2] For simplicity, we shall

consider here only the low-rate, expurgated portion of the bound, which involves a much simpler problem than the maximization required at higher rates. This restriction will still allow us to obtain the straight-line bound.

After a suitable normalization, the expurgated bound has the essential form

$$Pe \leq \exp -N[E_{xK}(\rho, \underline{p}, r) - \rho R] \tag{2}$$

$$E_{xK}(\rho, \underline{p}, r) = -\rho \ln \left[\int_0^\infty \int_0^\infty p(x)\, p(x_1)\, e^{r[f(x)+f(x_1)]}\, H(x, x_1)^{K/\rho}\, dx dx_1 \right] \tag{3}$$

$$H(x, x_1) = \frac{(1+x^2)^{1/2} \left(1+x_1^2\right)^{1/2}}{1 + \frac{1}{2}\left(x^2+x_1^2\right)} . \tag{4}$$

In these equations, $p(x)$ is the probability density on the input for any one channel use, $r \geq 0$, and $\rho \geq 1$. Also, $\int_0^\infty f(x)\, p(x)\, dx = 0$, where $f(x) = x^2 - \frac{A}{K}$, representing an input energy constraint (i.e., $A = \frac{1}{N} \frac{PT}{N_o}$, where $PT = \frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N} x_{nm}^2$ is the average input signal energy). For the expurgated bound, maximization of the exponent for a given K, ρ, A can be simplified by assuming $K = 1$ and replacing ρ by $\frac{\rho}{K} = \rho'$ and A by $\frac{A}{K} = A'$, and then maximizing, for if we let

$$E_{xK}(\rho, A) = \max_{p(x), r} E_{xK}(\rho, \underline{p}, r) \tag{5}$$

for a given A and ρ, then

$$E_{xK}(\rho, A) = K E_{x1}(\rho', A'). \tag{6}$$

This reduces the maximization to a two-parameter one, although now the range of ρ' is $\rho' > 0$ when $\rho \geq 1$, to provide for all values of K.

3. Optimization over p(x)

It is possible to derive sufficient conditions on $p(x)$ and r to maximize $E_{x1}(\rho, \underline{p}, r)$, subject to the energy constraint. For $0 < \rho < \infty$ the derivation depends on the fact that $H(x, x_1)^{1/\rho}$ is non-negative definite for the particular $H(x, x_1)$ considered here. The resulting sufficient condition is

$$\int_0^\infty p(x_1)\, e^{r[f(x)+f(x_1)]}\, H(x, x_1)^{1/\rho}\, dx_1 \geq \int_0^\infty \int_0^\infty p(x)\, p(x_1)\, e^{r[f(x)+f(x_1)]}\, H(x, x_1)^{1/\rho}\, dx dx_1 \tag{7}$$

for all x, with equality when $p(x) > 0$.

At zero rate, when $\rho = \infty$, the situation changes somewhat, for the problem now

becomes

$$\max_{p(x)} - \int_0^\infty \int_0^\infty p(x)\, p(x_1) \ln H(x, x_1)\, dx dx_1$$

with the same energy constraint as before, but now $\ln H(x, x_1)$ is not non-negative definite. All that the proof actually requires, however, is that $\ln H(x, x_1)$ be non-negative definite with respect to all functions that can be represented as the difference of two finite energy probability distributions, and it is possible to show that such is indeed the case. For $\rho = \infty$, the resulting sufficient condition for the maximization is

$$\int_0^\infty p(x_1) \ln H(x, x_1)\, dx_1 \geq \int_0^\infty \int_0^\infty p(x)\, p(x_1) \ln H(x, x_1)\, dx dx_1 + \lambda(x^2 - A) \qquad (8)$$

for all x, again with equality when $p(x) > 0$. Note that the conditions above have only been proved sufficient and not necessary. Hence it is conceivable that some other probability density exists which results in an exponent equal to the one obtained by a p(x) satisfying the conditions. There cannot, of course, be a p(x) that gives a larger exponent.

For a range of A and ρ, it is possible to show that a p(x) consisting of two impulses, one of which is at the origin, will satisfy the applicable sufficient condition. In particular, for $\rho = \infty$ and all A, such a probability density will satisfy conditions (8). For any ρ, $0 < \rho < \infty$, there is an $A_2 < \infty$ such that, for $A \leq A_2$, a two-impulse p(x) satisfies condition (7) and thus is optimum. Also, as $A \to 0$, this type of p(x) is optimum for all ρ, and the resulting exponent is the same as the infinite-bandwidth, orthogonal signal exponent obtained by Kennedy[3] (actually, this is true for the whole random coding bound, and not just the expurgated portion). When a two-impulse solution is optimum, it signifies that it is sometimes advantageous to conserve energy by not using an individual channel, so that when a channel is used, it is with an "optimum" value of energy-to-noise ratio per diversity path.

4. Zero-Rate Upper Bound

As we have mentioned, the case of zero rate is especially simple because the optimum probability density always consists of two impulses, regardless of the other system parameters. Also, holding $R = 0$ implies $\rho = \infty$ independent of other parameters, while if R exceeds zero, a change in system parameters will usually require a corresponding change in the value of ρ.

To illustrate the character of the results we consider a specific waveform set which, although often useful, is not always an efficient one. Let the basic signals have bandwidth W_s, and last for a time T_s, $T_s W_s = 1$. A guard space between signals of $B + 1/L$ cps in frequency and $L + 1/B$ sec in time should serve to make the signals

independent and orthogonal at the channel output, where B is the channel Doppler spread, and L is the multipath spread. Then if we consider a total signaling interval of T sec and a bandwidth of W cps, we can roughly express N and K in terms of the preceding parameters.[1] For example, with the choice of $BT_s > 1$, $LW_s < 1$,

$$K \cong BT_s \tag{9}$$

and

$$N \cong \frac{TWS}{(1+S+K)\left(1+S+\frac{S}{K}\right)}, \tag{10}$$

where S = BL is the channel-spread factor.

We want to choose K (and thereby N) to maximize the exponent $NE_{xK}(\infty, A)$. Since

$$\text{exponent} = \frac{PT}{N_o}\left\{\frac{E_{x1}\left(\infty, \frac{A}{K}\right)}{\frac{A}{K}}\right\}, \tag{11}$$

we see from Fig. XXVI-1 that to maximize the exponent, we must minimize $\frac{A}{K} = \frac{PT}{N_o}\frac{1}{NK}$, which requires making K as large as possible. From (9) we that this can be

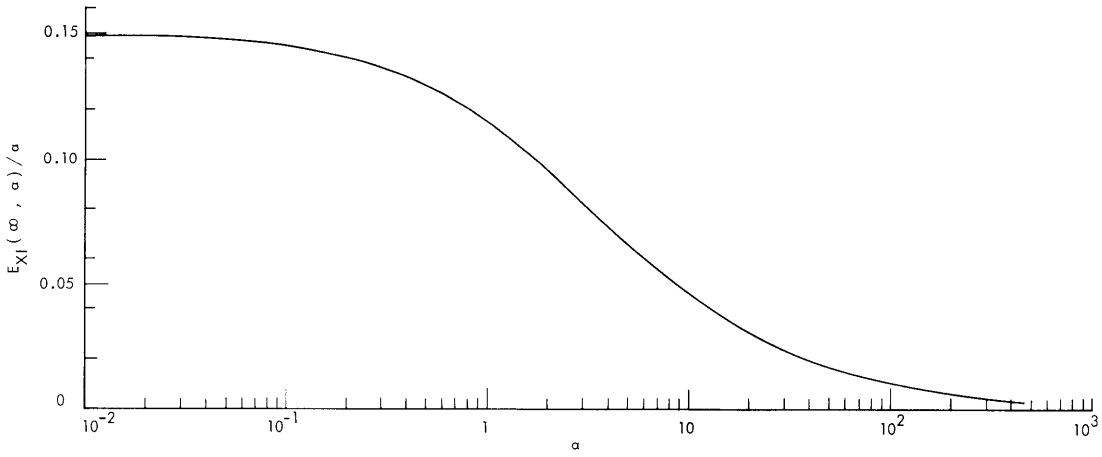

Fig. XXVI-1. $\frac{E_{x1}(\infty, \alpha)}{\alpha}$ vs α.

done by using very long basic signals, with a resulting value

$$\frac{A}{K} \cong \frac{1+\frac{1}{S}}{W}\left(\frac{P}{N_o}\right). \tag{12}$$

As we have noted, $W \to \infty$ means $A/K \to 0$ and $\frac{E_{x1}(\infty, A/K)}{A/K} \simeq .15$, which results in the infinite-bandwidth exponent found by Kennedy. Note that, for finite bandwidths, the exponent obtained here is still a function of W, even at zero rate. This contrasts with the case of the additive Gaussian noise channel, for which the zero-rate exponent is independent of bandwidth.

5. Rates Greater than Zero

For positive rates, the situation becomes more complex, and the best p(x) is still unknown. It has not been possible to show that the exponent is always maximized by a discrete p(x) (the proof for two impulses involves solving for the probabilities and positions, and this quickly becomes laborious for more than two), but it seems reasonable to believe that such is the case. If, for example, we let p(x) consist of a grid of impulses spaced along the x-axis, and numerically optimize on the impulse probabilities, we find that most probabilities are zero, and the nonzero ones are widely separated. As the grid spacing is reduced, the nonzero impulses change positions and probabilities slightly, and the others remain zero. If the distribution that is arrived at in this manner involved more and more impulses with smaller and smaller probabilities as the grid spacing was

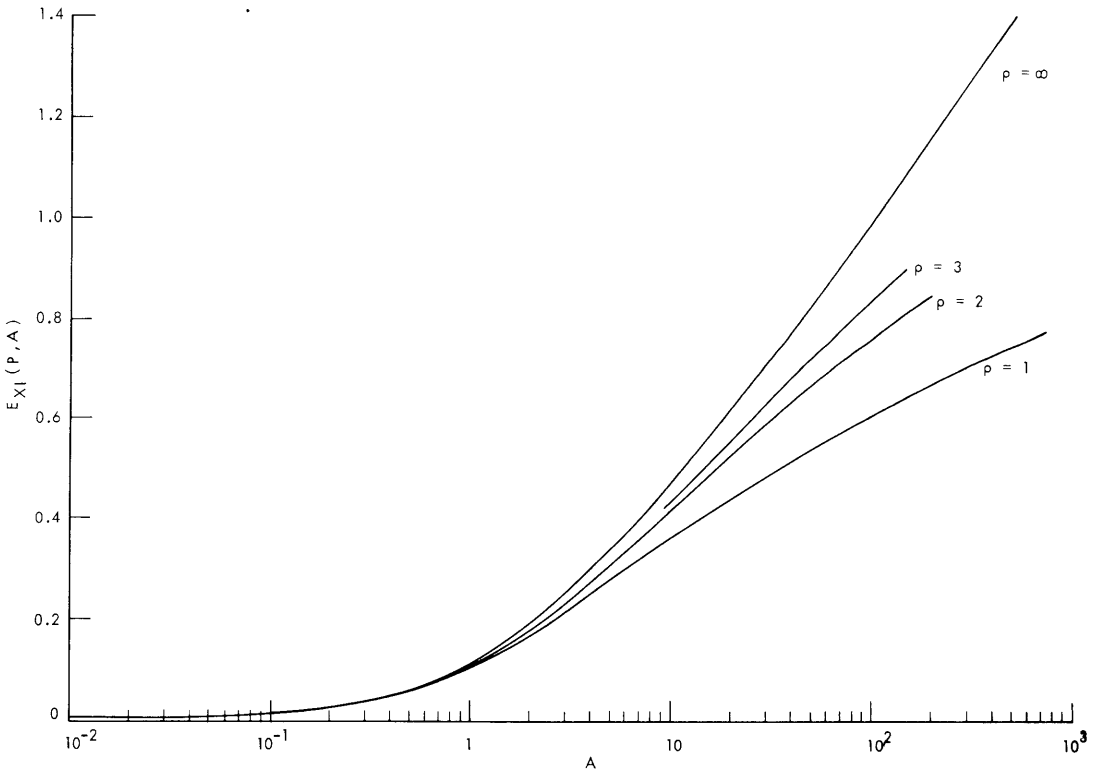

Fig. XXVI-2. $E_{x1}(\rho, A)$ vs A for several values of ρ.

reduced, we would say that it was beginning to approximate a continuous distribution. Instead, the only effect of reduced grid spacing is an apparent relocation of the impulses to better postiions, and the distribution still looks impulsive.

As A is increased from zero for a given value of ρ, the optimum probability distribution starts as two impulses, then appears to become three, then four, etc. This has been true for all values of ρ for which distributions have been computed.

Whether or not discrete solutions are always exactly optimum, the type of computation described above indicates that they are approximately optimum, and consequently allows computation of approximate results, such as those illustrated in Fig. XXVI-2.

If we consider one-channel use when $\rho = \infty$, with fixed energy but variable diversity, or, equivalently, fix N and A but allow K to vary, increasing K results in a monotonically increasing exponent, and $K \to \infty$ gives the infinite-bandwidth result. When $\rho < \infty$, it appears that we cannot get the infinte-bandwidth result merely by increasing K with N and A fixed, for there is a value of K beyond which the exponent starts decreasing again. Of course, allowing N to become large, for example, by increasing the bandwidth, will always lead to the infinite-bandwidth exponent.

J. S. Richters

References

1. R. S. Kennedy and I. L. Lebow, "Signal Design for Dispersive Channels," IEEE Spectrum, Vol. 1, No. 3, p. 231, March 1964.
2. R. G. Gallager, "A Simple Derivation of the Coding Theorem and Some Applications," IEEE Trans. Vol. IT-11, No. 1, pp. 3-18, January 1965.
3. R. S. Kennedy, Performance Limits for Fading Dispersive Channels (to be published by John Wiley and Sons, Inc., New York).

B. OPTICAL PROPAGATION THROUGH A TURBULENT ATMOSPHERE

It is known that atmospheric turbulence severely affects the performance of optical communication systems that involve propagation through the earth's atmosphere. Neither the fundamental limitations that this turbulence imposes upon the attainable communication reliability nor the signal and receiver structures that are most effective in combating it are known, however. To answer these questions it is necessary to determine the statistical properties of the received signal, or field, that exists over the receiving aperture.

The most casual examination convinces one that a rigorous, diffraction-theory treatment of the problem is unappealing, if not impossible. A further examination suggests that the most generally accepted technique of approximation is that due to Rytov,[1] as presented by Tatarski.[2] For point sources, plane waves, and "uncollimated

beams" this approach yields results that are in good agreement with experiment.[3] Recently, this technique has been applied to the near field of collimated beams by Tatarski,[4] but the complexity of the results limits their utility in a communication analysis. Moreover, the behavior of the far field is still unknown.

Another technique of approximation is provided by geometric optics.[5] Its utility in the study of atmospheric optical propagation has been in doubt because it has yielded results that fail to agree with experiment and because heuristic arguments have been advanced to suggest that diffraction effects are of basic importance in such studies.[6-8] Geometric optics does lead, however, to a relatively simple, and unified, model for the effects of turbulence in both the near and far fields. Such simplicity is almost as important to any subsequent communication analysis as is extreme precision. Moreover, the arguments suggesting that diffraction effects are important only consider the behavior of small isolated elements of the atmosphere and ignore the averaging, or masking, of diffraction effects that results from the superposition of numerous such elements. Therefore, a study of atmospheric optical propagation predicated upon geometric optics was initiated. The preliminary results suggest that the lack of agreement between theory and experiment has resulted from a failure to distinguish between rays and their mean values and between points on rays and points in space. This has led us to pursue the geometrical optics approach beyond the level required for communication analysis.

1. Atmospheric Model

The random space-time variations of the atmospheric refractive index cause the statistical fluctuations observed in optical propagation. These variations can be reasonably treated as locally isotropic and, for simplicity, one also often supposes that they are homogeneous. This latter supposition is not crucial, and can be removed at the cost of complicating the results. We shall limit our discussion here to the behavior of the field at a single instant of time. Temporal problems will be treated subsequently.

The only knowledge that we require of the refractive-index variations is that they be quite small, the variation at points separated by distances of the order of meters be independent, and the mean and correlation function, or the structure function, of the variations be given. The most realistic choice for a structure function is probably that obtained from the Obukhov-Kolmogorov theory of turbulence.[9] Since this choice leads to rather complicated expressions, however, we shall often consider other simpler functions to illustrate the results. Reiger[10] has developed an approximation to the turbulence spectrum in the inertial subrange which is useful in many calculations because of its form. With the Reiger spectrum the calculations are reduced to those required for a Gaussian-shaped correlation function.

2. Geometric Optics

It suffices to consider the sinusoidal steady state. The complex fields are then of the form

$$E(\vec{r}) = \vec{e}(\vec{r}) \exp\left[j\frac{2\pi}{\lambda} \mathscr{L}(\vec{r}) \right] \tag{1a}$$

$$H(\vec{r}) = \vec{h}(\vec{r}) \exp\left[j\frac{2\pi}{\lambda} \mathscr{L}(\vec{r}) \right]. \tag{1b}$$

By the approximation of geometric optics, $\mathscr{L}(\vec{r})$, satisfies the Eikonal equation

$$|\nabla \mathscr{L}(\vec{r})|^2 = n^2(\vec{r}), \tag{2}$$

where $n(\vec{r})$ is the refractive index as a function of position, $\vec{r}$. In this expression we have employed the notion of a frozen atmosphere and have suppressed the time variation of the refractive index. The quantity $\mathscr{L}(\vec{r})$ is the optical path length, while $\frac{2\pi}{\lambda}\mathscr{L}(\vec{r})$ is the phase function.

For our purposes, the utility of geometric optics lies in the ray picture that it provides. These rays are defined to be the set of trajectories perpendicular to the constant phase surfaces ($\mathscr{L}(r)$ = constant). Each ray can be specified by the variation of its position vector, $\vec{r}_\tau$, with the parameter τ. The parameter τ specifies position along the ray and is roughly proportional to arc length ℓ as $d\ell = n d\tau$. It is also convenient to introduce the ray direction, $\vec{u}_\tau$, which is the derivative of $\vec{r}_\tau$ with respect to τ. These vectors are related to each other and to the optical phase length by the expressions

$$\vec{u}_\tau = \frac{d\vec{r}_\tau}{d\tau} = \nabla \mathscr{L}(\vec{r}_\tau) \tag{3a}$$

and

$$\frac{d\vec{u}_\tau}{d\tau} = \frac{\nabla n^2(\vec{r}_\tau)}{2}. \tag{3b}$$

Integration of these equations yields

$$\vec{u}_\tau = \vec{u}_o + \frac{1}{2}\int_0^\tau \left[\nabla n^2(\vec{r}_\sigma)\right] d\sigma \tag{4a}$$

and

$$\vec{r}_\tau = \vec{r}_o + \vec{u}_o\tau + \frac{1}{2}\int_0^\tau (\tau-\sigma)\left[\nabla n^2(\vec{r}_\sigma)\right] d\sigma, \tag{4b}$$

where the quantities $\vec{u}_o$ and $\vec{r}_o$, the initial ray direction and position at the source of radiation, serve to specify the ray in question.

The field quantities $\mathscr{L}(\vec{r})$, $\vec{e}(\vec{r})$, and $\vec{h}(\vec{r})$ also can be evaluated by integrals along the rays. In particular, the variation of the phase along any given ray is governed by the differential equation

$$\frac{d\,\mathscr{L}(\vec{r}_\tau)}{d\tau} = n^2(\vec{r}_\tau) \tag{5a}$$

which yields

$$\mathscr{L}(\vec{r}_\tau) = \mathscr{L}_o + \int_0^\tau n^2(\vec{r}_\sigma)\, d\sigma, \tag{5b}$$

where $\mathscr{L}_o$ is the initial phase. For simplicity, $\mathscr{L}(\vec{r}_\tau)$ will often be written as $\mathscr{L}(\tau)$. Also,

$$\frac{|\vec{e}(\vec{r}_\tau)|^2}{|\vec{e}_o|^2} = \exp\left[-\int_0^\tau \left[\nabla^2 \mathscr{L}(\vec{r}_\sigma)\right] \frac{d_\sigma}{n(\vec{r}_\sigma)}\right] \tag{6a}$$

and

$$\frac{|\vec{h}(\vec{r}_\tau)|^2}{|\vec{h}_o|^2} = \frac{|\vec{e}(\vec{r}_\tau)|^2}{|\vec{e}_o|^2}, \tag{6b}$$

where $\vec{e}_o$ and $\vec{h}_o$, are the initial vlaues. In Eq. 6 we have approximated $\left[\frac{n(\vec{r}_\tau)}{n(\vec{r}_o)}\right]^2$ by unity.

To avoid the necessity of evaluating the Laplacian of $\mathscr{L}(\vec{r}_\tau)$, it is sometimes desirable to employ the approximation

$$\nabla^2 \mathscr{L}(\vec{r}_\sigma) \approx \frac{\partial \ln n(\vec{r}_\sigma)}{\partial\sigma} + \int_0^\sigma \nabla^2 n^2(\vec{r}_\rho)\, d\rho. \tag{7}$$

Upon introducing this into Eq. 6a, evaluating two integrals, and again approximating $\frac{n(\vec{r}_\tau)}{n(\vec{r}_o)}$ by unity we obtain

$$\frac{|\vec{e}(\vec{r}_\tau)|^2}{|\vec{e}_o|^2} = \frac{|\vec{h}(\vec{r}_\tau)|^2}{|\vec{h}_o|^2} \approx \exp\left[-\int_0^\tau (\tau-\sigma)\left[\nabla^2 n^2(\vec{r}_\sigma)\right] d\sigma\right]. \tag{8}$$

To complete the specification of the fields, it is necessary to specify their directions.

By the approximations of geometric optics, the $\vec{e}$ and $\vec{h}$ vectors are perpendicular to each other and to the ray direction vector. Thus it suffices to determine the angle, θ, between the direction of $\vec{e}_o$ and of $\vec{e}(\vec{r}_\tau)$, i.e., the depolarization angle. Although a precise evaluation of this angle is difficult, it can be shown that it is essentially equal to the change in the ray direction vector.[11] That is, there is no depolarization as such but only the rotation of $\vec{e}(\vec{r}_\tau)$ that is associated with the changing direction of $\vec{u}(\tau)$. In particular, subject to the suppostion that θ is not too large, it is easy to show that

$$\theta_\tau \approx \frac{|\vec{e}_o \cdot \vec{u}_\tau|}{|\vec{e}_o|^2}. \tag{9}$$

Equations 4 through 9, in conjunction with the Central Limit theorem, lead to a complete statistical description of the field at any point, τ, on a ray. Specifically, for any given value of τ that is sufficiently large $\vec{r}_\tau$, $\vec{u}_\tau$, $\mathscr{L}(\vec{r}_\tau)$ and $\ln\left[|\vec{e}(\vec{r}_\tau)|^2\right]$ will be Gaussian random (vector) variables. More generally, it is reasonable to suppose that the values of these quantities at different points on the same ray, or on different rays, are jointly Gaussian. Finally, the remaining quantities $\vec{h}(\vec{r}_\tau)$ and θ_τ are determined by Eqs. 6b and 9.

3. Moments

Since all of the field quantities are determined by a set of Gaussian random variables, it suffices to know their means and covariances. To evaluate these quantities, we invoke an approximation that is best explained by an example, By virtue of Eq. 4b, the expected value of $\vec{r}_\tau$ at the parameter point τ is

$$E[\vec{r}_\tau] = \vec{r}_o + \tau\vec{u}_o + \frac{1}{2}\int_0^\tau (\tau-\sigma)\, E_{\vec{r}}\left[\nabla E_{n|\vec{r}}\left[n^2(\vec{r}_\sigma)\right]\right] d\sigma, \tag{10}$$

where $E_{n|\vec{r}}$ denotes the conditional expected value of $n^2(\vec{r}_\sigma)$ at the parameter point σ on the ray, given that the position at that point is $\vec{r}_\sigma$.

We next claim that the conditional average is approximately equal to the unconditional average evaluated at the (random) point $\vec{r}_\sigma$. This is so because $\vec{r}_\sigma$ is controlled by the large-scale behavior of the atmosphere, while $n(\vec{r})$ is a local quantity: hence, knowledge of $\vec{r}_\sigma$ yields little information about $n(\vec{r}_\sigma)$. Thus we obtain

$$E[\vec{r}_\tau] \approx \vec{r}_o + \tau\vec{u}_o + \frac{1}{2}\int_0^\tau (\tau-\sigma)\, E_{\vec{r}_\sigma}\left[\nabla E_n\left[n^2(\vec{r}_\sigma)\right]\right] d\sigma, \tag{11}$$

where E_n is the unconditional average of n^2 evaluated at the random position $\vec{r}_\sigma$, and

$E_{\vec{r}_\sigma}$ denotes the average of the indicated quantity with respect to the ray position vector, $\vec{r}_\sigma$. The approximation that leads from Eq. 10 to Eq. 11 can be used to evaluate all of the required averages.

For example, the approximation, in conjunction with the supposition that the refractive index is homogeneous in space, leads one to conclude that

$$E[\vec{r}_\tau] = \vec{r}_o + \tau\vec{u}_o \tag{12a}$$

$$E[\vec{u}_\tau] = \vec{u}_o \tag{12b}$$

$$E[\mathscr{L}(\vec{r}_\tau)] = \mathscr{L}_o + \tau \tag{12c}$$

and

$$E\left[\ln \frac{|\vec{e}(\vec{r}_\tau)|^2}{|\vec{e}(o)|^2}\right] = 0. \tag{12d}$$

Here, we have supposed that the unconditional average of $n^2(\vec{r})$ is unity.

The covariance of the quantities above can be obtained by straightforward, but laborious, calculation. Since the results are cumbersome, we shall only discuss the covariance of the position vector along a single ray. This quantity is chosen because it involves more difficulties than do the other covariances.

The covariance of the i^{th} component of the position vector at the parameter value τ and the j^{th} component at parameter value τ' is easily shown to be

$$C_{y_i y_j}(\tau, \tau') = E[y_i(\tau)y_j(\tau')]$$

$$\approx -\frac{1}{2}\int_0^\tau \int_0^{\tau'} (\tau-\sigma)(\tau'-\sigma')\, E_{\vec{r}_\sigma, \vec{r}_{\sigma'}}\left[\frac{\partial^2}{\partial q_i \partial q_j'} D_n(\vec{r}_\sigma, \vec{r}_{\sigma'})\right] d\sigma d\sigma', \tag{13a}$$

where $\vec{y}(\tau)$ is the zero mean position vector defined as

$$\vec{y}(\tau) = \vec{r}_\tau - E(\vec{r}_\tau)$$

$$\vec{y}(\tau) = \sum_{k=1}^{3} y_k \vec{i}_k, \tag{13b}$$

q_i and q_j are the variables associated with the i^{th} and j^{th} rectangluar coordinates, respectively, and the refractive index structure function is defined by the relation

$$D_n(\vec{r}_\sigma, \vec{r}_{\sigma'}) = E_n\left\{[n(\vec{r}_\sigma) - n(\vec{r}_{\sigma'})]^2\right\}. \tag{13c}$$

In the derivation of Eq. 13a the assumptions that the refractive index is homogeneous and that the variations are small have been exploited to approximate

$$\frac{\partial^2}{\partial q_i \partial q_j'} E\left[n^2(\vec{r}_\sigma) n^2(\vec{r}_{\sigma'}) \right]$$

by

$$-2 \frac{\partial^2}{\partial q_i \partial q_j'} D_n(\vec{r}_\sigma, \vec{r}_{\sigma'}).$$

Note that the covariance of the ray position enters into both sides of Eq. 13a. Thus it provides an equation which the covariance must satisfy, rather than an expression from which it can be directly evaluated. A direct solution of this equation appears to be quite difficult although it may be possible in some special situations.

An alternative, and commonly employed, approach is to replace the position vectors in the integrand by their average values so as to avoid the expectation operation with respect to the position vectors. That is, one ignores the ray motion and evaluates the integrals along the average, or unperturbed, ray to obtain a first-order estimate of the covariance. If desired, the resulting expression for the ray covariance can be used to obtain a second-order estimate: more generally, one can iterate the scheme indefinitely. We have not yet established, however, that the higher order estimates converge to the true covariance function.

Once the ray position vector covariance has been determined, the other covariances can be obtained by straightforward calculations. With the exception of the covariance for the ray direction vector, which is required later, the covariances of the other quantities will be omitted in the interest of brevity. The covariance of the i^{th} and j^{th} components of the ray direction vector, $\vec{u}_\tau$, is

$$\begin{aligned} C_{\mu_i \mu_j}(\tau, \tau') &= E[\mu_i(\tau)\mu_j(\tau')] \\ &= -\frac{1}{2} \int_0^\tau \int_0^{\tau'} E_{\vec{r}_\sigma, \vec{r}_{\sigma'}} \frac{\partial^2}{\partial q_i \partial q_j'} D_n(\vec{r}_\sigma, \vec{r}_{\sigma'}) \, d\sigma d\sigma', \end{aligned} \tag{14a}$$

where $\vec{\mu}(\tau)$ is the zero-mean ray direction vector defined as

$$\begin{aligned} \vec{\mu}(\tau) &= \vec{u}_\tau - E[\vec{u}_\tau] \\ \vec{\mu}(\tau) &= \sum_{k=1}^{3} \mu_k \vec{i}_k. \end{aligned} \tag{14b}$$

Certain conclusions can be drawn from Eqs. 13a and 14a without explicity solving them. One conclusion, which is needed later, is that the components of $\vec{y}(\tau)$ are uncorrelated and also that the components of $\vec{\mu}(\tau)$ are uncorrelated. Also, we can conclude that the two components of $\vec{y}(\tau)$ perpendicular to the direction of propagation have the same variance. A similar statement is also true for $\vec{\mu}(\tau)$.

Rather than pursue the moment calculations further, we now consider the distinction between the quantities defined on rays and those measured in space. This distinction will then be illustrated by considering the phase covariance function.

4. Ray-to-Space Transformations

The results stated above do not directly specify the field at any fixed point in space; rather, they specify the field values at points on rays and also the statistical behavior of the ray spatial position. Thus, to determine the behavior of the field at a point in space, one must, in principle, determine the probability that any particular ray passes through the point in question and then determine the conditional statistics of the field on that ray. This ray-to-space transformation appears to have been overlooked in previous applications of geometric optics.

In the ray-to-space transformation, the question of how many rays may go through a given point in space arises. This is essentially the question of whether interference effects are important or can be neglected. It seems clear that interference effects must be accounted for in some situations, e.g., converging beams. The large coherent bandwidths observed in optical propagation through the atmosphere suggest, however, that interference is not a first-order effect for unfocused beams over path lengths of several kilometers.[12,13] In the sequel interference will be neglected, and at most a single ray will be assumed to go through any given point.

Although the ray-to-space transformation has not yet been completed for all of the field quantities, the preliminary results suggest that it will remedy many of the defects normally attributed to geometrical optics, and permit the solution of some previously unsolved problems. To illustrate the possiblities, we shall now examine the phase covariance in more detail.

5. Phase Covariance

In this section the source distribution is assumed to be a uniform plane wave propagating in the z-direction. Thus all rays have the same initial direction vector and phase value. As usual, homogeneity and local isotropy of the refractive-index variations are assumed.

A calculation of the phase covariance for two points lying on a line parallel to the direction vectors at the source provides an illustration of the importance of accounting for ray motion. The solution presented here is only a first-order solution because most

integrations will be carried out along unperturbed ray paths. The ray parameter τ is thus assumed to be the same for any two points in a plane perpendicular to the source direction vectors. Furthermore, as we have mentioned, it is assumed that only one ray goes through each point in space.

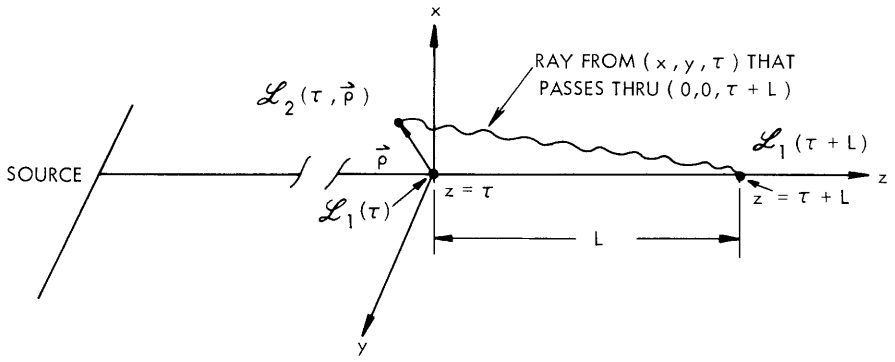

Fig. XXVI-3. Coordinate frame for calculation of phase covariance.

The situation of interest is shown in Fig. XXVI-3 where the coordinate system is indicated. The quantity to be calculated is $E[(\mathscr{L}_1(\tau) - \overline{\mathscr{L}_1(\tau)})(\mathscr{L}_1(\tau+L) - \overline{\mathscr{L}_1(\tau+L)})]$, the joint central moment of the phase at the intersection of the z-axis and the planes $z = \tau$ and $z = \tau + L$. Note that, because of the ray motion, the ray that passes through the last point may not pass through the former point. Instead it passes through another point in the plane $z = \tau$ whose position is indicated by the vector $\vec{\rho} = x\vec{i}_x + y\vec{i}_y$. The phase at this second point in the plane $z = \tau$ is denoted by $\mathscr{L}_2(\tau;\vec{\rho})$.

To simplify the notation, let

$$S_1 = \mathscr{L}_1(\tau) - \overline{\mathscr{L}_1(\tau)} \tag{15a}$$

$$S_2 = \mathscr{L}_2(\tau+L) - \overline{\mathscr{L}_1(\tau+L)} \tag{15b}$$

$$S_3(\vec{\rho}) = \mathscr{L}_2(\tau;\vec{\rho}) - \mathscr{L}_2(\tau;\vec{\rho}) \tag{15c}$$

The quantity to be calculated now is $E[S_1S_2]$. It is convenient to first compute the conditional average, given that the ray passing through $(0, 0, \tau+L)$ has the position vector $\vec{\rho}$ in the plane $z = \tau$, and then to average over $\vec{\rho}$. Thus

$$\begin{aligned} E[S_1S_2|\vec{\rho}] &= E[S_1(S_3(\vec{\rho}) + \Phi(\vec{\rho}))] \\ &= E[S_1S_3(\vec{\rho})] + E[S_1\Phi(\vec{\rho})], \end{aligned} \tag{16}$$

where $\Phi(\vec{\rho})$ is the zero-mean increment in the phase as the ray propagates from (x, y, τ) to $(0, 0, \tau + L)$. If L is large relative to the correlation distance of the refractive-index variations, then the second expectation on the right can reasonably be assumed to be approximately zero. The conditional expectation then becomes

$$E(S_1 S_2 | \vec{\rho}) = C_s(\rho; \tau), \tag{17}$$

where $C_s(\rho; \tau)$ is the spatial phase covariance function in a plane perpendicular to the direction of propagation.

To remove the conditioning in Eq. 17, it is necessary to average the conditional expectation over $\vec{\rho}$. To do this exactly, the probability must be used that the ray from (x, y, τ) goes through $(0, 0, \tau+L)$ and that all other rays do not. A much simpler approximate method is to think of replacing the ray through $(0, 0, \tau+L)$ with a pseudoray directed from the point $(0, 0, \tau+L)$ back toward the plane $z = \tau$. This pseudoray is assumed to have a direction vector which is the negative of the direction vector of the ray that goes through $(0, 0, \tau+L)$.

It is now possible to use this pseudoray to calculate the probability that the actual ray emanated from various parts of the $z = \tau$ plane and thus to carry out the desired averaging. Since the direction vector of the actual ray is also random, one must average over its value. This average involves the density of the direction vector components at a point in space which, in turn, requires a ray-to-space transformation. In the following discussion this ray-to-space transformation will be neglected, and it will be assumed that the form of the density is not changed and, furthermore, that the variance associated with the direction vector components for a ray parameter value of $\tau + L$ can be used. We believe that these assumptions do not significantly alter the results.

As we have noted, the ray position vector components are independent Gaussian random variables. The ray direction vector components are also independent and Gaussian. Thus, the conditional density of the x component of the pseudoray position is

$$p_{X|u_x}(X) = \frac{1}{\sqrt{2\pi}\,\sigma_x} \exp\left[-\frac{1}{2}\,\frac{(X - u_x L)^2}{\sigma_x^2}\right], \tag{18}$$

where u_x is the x component of the pseudoray direction vector. As u_x has a zero mean, its density is

$$p_{u_x}(U) = \frac{1}{\sqrt{2\pi}\,\sigma_{u_x}} \exp\left[-\frac{1}{2}\,\frac{U^2}{\sigma_{u_x}^2}\right]. \tag{19}$$

Similar equations can be written for the y component of the pseudoray position and direction vectors. It is thus possible, by using the independence, to obtain the

unconditional joint density for the two lateral ray position coordinates, x and y. Finally, the density of $\rho = |\vec{\rho}| = \sqrt{x^2 + y^2}$ is given by

$$p_\rho(P) = \frac{P}{\left(L^2\sigma_{u_x}^2 + \sigma_x^2\right)} \exp\left[-\frac{1}{2}\frac{P^2}{L^2\sigma_{u_x}^2 + \sigma_x^2}\right], \tag{20}$$

where the facts that $\sigma_x^2 = \sigma_y^2$ and $\sigma_{u_x}^2 = \sigma_{u_y}^2$ have been used.

Before this density can be used to complete the determination of $E[S_1S_2]$ it is necessary to determine the lateral spatial covariance function of the phase. The first-order solution (integration along the unperturbed rays) is

$$C_s(\rho;\tau) \approx 4\int_0^\tau \int_0^\tau C_n(|\vec{r}_\sigma - \vec{r}_{\sigma'}|)\, d\sigma d\sigma', \tag{21}$$

where $C_n(|\vec{r}_\sigma - \vec{r}_{\sigma'}|)$ is the refractive index covariance function, and the integration is along two parallel paths separated by a distance, ρ. The refractive index can be written

$$n = 1 + \delta, \tag{22}$$

where δ is very small, and the approximation exists because terms in δ^3 and δ^4 have been dropped.

Under the assumption of a Gaussian covariance function for the refractive-index variations,

$$C_n(\rho) = \overline{\delta^2} \exp\left[-\frac{\rho^2}{a^2}\right], \tag{23}$$

the equation above reduces to

$$C_s(\rho;\tau) = 4\sqrt{\pi}\ \overline{\delta^2}\ a\ \tau \exp\left(-\frac{\rho^2}{a^2}\right). \tag{24}$$

With the Gaussian covariance function of Eq. 23, and a first-order solution neglecting terms in δ^3 and δ^4, the variance of a lateral component of the ray direction vector and a lateral component of the ray position vector can be written

$$\sigma_{u_x}^2(\tau) = 2\sqrt{\pi}\,\frac{\overline{\delta^2}}{a}\,\tau \tag{25}$$

$$\sigma_x^2(\tau) = \frac{2}{3}\sqrt{\pi}\,\frac{\overline{\delta^2}}{a}\,\tau^3. \tag{26}$$

It is now possible to calculate the desired expectation. Using the density of Eq. 20

to remove the conditioning in Eq. 17, where the value of $C_s(\rho;\tau)$ is given by Eq. 21, one obtains

$$E[S_1S_2] = \frac{4\sqrt{\pi}\,\overline{\delta^2}\,a\,\tau}{1+\frac{2\left[L^2\sigma_{u_x}^2(\tau+L)+\sigma_x^2(L)\right]}{a^2}}, \tag{27}$$

and by using the results of Eqs. 25 and 26 this reduces to

$$E[S_1S_2] = \frac{4\sqrt{\pi}\,\overline{\delta^2}\,a\,\tau}{1+\frac{2\,\overline{\delta^2}\,\sqrt{\pi}}{a^3}\left(2L^2\tau+\frac{8}{3}L^3\right)}. \tag{28}$$

In order to provide a result for comparison, we shall now consider the phase convariance, using unperturbed rays with no ray-to-space transformation. The equation of interest,

$$E[S_1S_2] = 4\int_0^{\tau+L}\int_0^{\tau} C_n(|\vec{r}_\sigma-\vec{r}_{\sigma'}|)\,d\sigma d\sigma', \tag{29}$$

is readily obtained by modifying Eq. 21. The integration in Eq. 29 is along the z-axis in Fig. XXVI-3. For the assumption of the Gaussian covariance function of Eq. 23 the result is

$$E[S_1S_2] \approx 2\overline{\delta^2}\left[(2\tau+L)\sqrt{\pi}\,a - L\,\mathrm{Erf}\left(\frac{\sqrt{2L}}{a}\right)\right], \tag{30}$$

where $\mathrm{Erf}(\cdot)$ is the error function, and it is assumed that $\tau \gg a$. Note that this phase covariance does not decrease as L increases but rather assumes a value equal to the phase variance at the point closest to the source.

The Gaussian covariance function of Eq. 23 is not a good model of the covariance of the refractive-index variations, but it does facilitate comparison of the results that are obtained with and without the ray-to-space transformation. As L becomes large relative to a, Eq. 30 reduces to the numerator of Eq. 28, and the difference in functional form is evident. Moreover, the extension to more realistic convariances, although cumbersome, is straightforward.

R. S. Kennedy, E. V. Hoversten

References

1. S. M. Rytov, "Diffraction of Light by Ultrasonic Waves," Izv. Akad. Nauk SSSR, Ser. Fiz., No. 2, 223 (1937).

2. V. I. Tatarski, Wave Propagation in a Turbulent Medium (translated by R. A. Silverman)(McGraw-Hill Book Company, New York, 1961), Chap. 7.

3. Ibid., Chap. 12.

4. A. I. Kon and V. I. Tatarski, "Fluctuation of the Parameter of a Spatially Bounded Light Beam in a Turbulent Atmosphere," Izvest. Vyss. Uch. Zav., Ser. Radiophysika 8, 870 (1965).

5. M. Born and E. Wolf, Principles of Optics (Pergamon Press, New York, 1959), Chap. III.

6. V. I. Tatarski, op. cit., p. 120 and Chap. 12.

7. L. A. Chernov, Wave Propagation in a Random Medium (translated by R. A. Silverman)(McGraw-Hill Book Company, New York, 1960), Chap. II.

8. G. E. Meyers, D. L. Fried, and M. P. Keister, Jr., "Experimental Measurements of the Character of Intensity Fluctuations of a Laser Beam Propagating in the Atmosphere," Technical Memorandum No. 252, IS-TM-651146, Electro-Optical Laboratory, Space and Information Systems Division, North American Aviation, Inc., September 1965.

9. J. L. Lumley and H. A. Panofsky, The Structure of Atmospheric Turbulence (Interscience Publishers, New York, 1964).

10. S. H. Reiger, "Starlight Scintillation and Atmospheric Turbulence," Astron. J. 68, 395 (1963).

11. A. A. M. Saleh, "Laser Wave Depolarization by Atmospheric Transmission," S.M. Thesis, Department of Electrical Engineering, M.I.T., Febuary 1967.

12. A. H. Mikesell, A. A. Hoag, and J. S. Hall, "The Scintillation of Starlight," J. Opt. Soc. Am. 41, 689 (1951).

13. E. G. Chatterton, "Optical Communications Employing Semiconductor Lasers," Technical Report 392, Lincoln Laboratory, M.I.T., June 1965.